SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER Final Report	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Task Influences in the Analytic-Intuitive Approach to Decision Making		5. TYPE OF REPORT & PERIOD COVERED 10/1/81 - 12/31/84
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) William C. Howell		8. CONTRACT OR GRANT NUMBER(s) N00014-82-C-0001
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Psychology Box 1892 Rice University Houston, TX 77251		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR197-074
11. CONTROLLING OFFICE NAME AND ADDRESS		12. REPORT DATE 12/22/84
		13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

task effects, uncertainty judgment, display effects, decision processes, information integration, predecision processes, decision heuristics, frequency estimation, optional stopping

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

see Abstract

DD FORM 1 JAN 73 1473

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

TASK INFLUENCES IN THE ANALYTIC-INTUITIVE APPROACH TO DECISION MAKING

William C. Howell

Rice University

Final Report

December 1984

This research was supported by the Engineering Psychology Programs, Office of Naval Research, ONR Contract N00014-82-C-001 Work Unit NR197-074.

CONTENTS

Task Influences in the Analytic-Intuitive Approach
to Decision Making

Final Report

William C. Howell

Rice University

Abstract

Three years of research have explored effects of some 14 formally irrelevant task variables on human judgment and decision behavior within the context of 5 different task scenarios. The variables (which represent input, processing, response, and feedback manipulations) and the scenarios (which represent a broad sampling of generic problem types) define a rather extensive domain within which it was possible to test implications of Hammond's "Cognitive Continuum Theory." Results were generally consistent with this conceptualization, although many of the specific generalizations seem to have practical and theoretical significance in their own right. A few of the more prominent conclusions were as follows:

1. Judgment is affected in both qualitative and quantitative ways by the manner in which information is displayed. Graphic coding promotes more "holistic" processing than does numeric coding, and consequently, better decisions under "pressure" situations. Acquisition of rules, however, is an inherently "analytic" process and is better served by numeric coding.

2. The "availability" heuristic can be induced by enhancing specific stimuli in an event stream.

3. Requiring overt judgments prior to a choice consistently improves the quality of that decision. In fact, such manual "pre-processing" compared very favorably with machine aiding in the multi-stage judgment problem.

4. Asymptotic judgment performance based on explicit rules does not require feedback for maintenance; in fact, it deteriorates with outcome feedback.

5. Time pressure, amount of information, rule availability, and rule complexity do not produce the "across the board" effects suggested by continuum theory. Each affects some aspect of performance on some task scenario, but not in the coherent fashion one would expect if the subject's cognitive approach were merely shifted along a unitary continuum. More needs to be learned about these effects--particularly as they interact.

INTRODUCTION

I. Overview of the Project

The broad purpose of the research conducted under Project NR197-074 was to clarify the effects of certain formally irrelevant task characteristics on judgment and decision performance. The primary integrating theme was Hammond's (1980, 1981) "cognitive continuum theory," which classifies a number of general task parameters with respect to their hypothesized influence on human thought. Some conditions, it is suggested, induce a more analytic (rule-based) approach; others, a more intuitive one.

Despite measurement problems which prevent a rigorous test of the theory, its logical and taxonomic features have proven useful in organizing the search for--and attack on--task influences. Variables were selected and manipulated in accordance with their expected cognitive implications, even though it is impossible to verify their precise location on the analytic-intuitive "continuum". To the extent that judgment or decision behavior changes in the predicted ways, something is learned about these particular task variables, and the basic tenents of the theory are strengthened. While extreme cases -- situations in which people behave relatively normatively (analytic extreme) or heuristically (intuitive extreme)--are not hard to find and have been well researched as discussed below, our focus has been on the middle ground--the area where subtle manipulations might have an important effect.

EXPERIMENTATION

In the course of this work, 15 formal experiments have been completed addressing a total of 14 variables within the context of five principal "task scenarios". This complete domain is represented in Figure 1 together with the particular experiments conducted under each variable-task combination.

A number of useful findings have emerged which are discussed in detail in 8 Technical Reports, two doctoral dissertations, two masters theses, and four manuscripts which have been or are being submitted to archival publications (one has appeared; two are under review; one is in preparation). These reports are listed at the end of the present section. Before examining the specific data produced by this work, it may be useful to discuss briefly the tasks, variables, and findings represented in Figure 1.

Task Scenarios. Of the five scenarios, two (A and B) were developed, pilot tested, and refined entirely in our laboratory. The third (C) is a standard vehicle for studying the integrative judgment process, the fourth (D) is a military adaptation of that problem which preserves its formal properties while expanding its research potential, and the fifth (E) is constructed from elements of (A) and (D) in combination. All have been programmed on our laboratory computers, although versions of (C) and (D) have been carried out in paper-and-pencil form to permit group administration. (It should be noted in this regard that because of their more complex nature and the individual administration requirement, tasks A, B, and D generally dictate much longer studies -- often requiring an entire semester-- than the group-administered, "one-shot" exercises typical of laboratory research). The five tasks are described briefly as follows.

A. Emergency resource allocation. Developed as a vehicle for studying the formation of and reactions to impressions of event

TASK SCENARIOS

MANIPULATED VARIABLES	A Emergency Resource Allocation	B Optional Stopping	C Personnel Decision (Information Integration)	D Threat Evaluation (Information Integration)	E Two-Stage Judgment
<u>Input</u>					
1. time constraint	4	11	2	3,13	6
2. information quantity			1	3, 9	
3. information quality			12		
4. stimulus enhancement	14, 15				
5. enhancement mode	15				
6. display order				9	
7. display format		10, 11	7	9	
<u>Processing</u>					
8. rule availability	4		1		
9. rule complex.			2	9, 13	
10. preprocessing (aiding)		10			5, 6
<u>Response</u>					
11. estimation requirement	14				5, 6, 8
12. resp. mode			7		8
<u>Feedback</u>					
13. type			12		
14. precision			12		

<u>Figure 1</u>. A summary of the variable-task domain defined and investigated in the project's first three years. Numbers merely designate specific experiments, in temporal sequence, that were completed; their location describes the content of the experiments.

uncertainty in a realistic setting, the problem consists of allocating resources (emergency units) in response to fire, police, and ambulance calls for a hypothetical city. The subject must acquire knowledge of tendencies, relations, pattern etc. associated with the flow of events in order to perform well. A variety of measures, each indicative of a different aspect of judgment/decision can be obtained (e.g. frequency and probability estimation of various kinds, predictions, choices, etc.) through probes administered following acquisition.

B. Optional stopping. Representing a situation common to many "real world" decision problems, this scenario involves trading time (and information) for decision quality. The subject chooses both when to stop sampling information, and which action to take at that point. This general type of problem has also been called the "information purchase" or "sequential decision" paradigm but, because of its complexity and time-consuming character, has not been popular in research. Our version consists of hurricane tracking; and the decision options, of evacuating or reinforcing a defined target area, or simply waiting. A realistic and sensitive cost/payoff system has been devised in terms of a lives-lost/saved index.

C. Personnel decision. This is a standard paradigm used in "policy-capturing" and "multiple-cue probability-learning" research (depending on whether the focus is on discovering established subjective values or plotting the acquisition of objectively specified values). Based on the Brunswik "lens model" (Brunswik, 1956; Hammond & Summers, 1972), it consists of a set of predictive items (cue values) that are stochastically related to a set of criterion values or states. The

subject judges criterion values for various sets (usually orthogonal combinations) of presented cue values, and performance is evaluated in terms of various "outcome" and "process" measures derived from the model (primarily involving linear regression weights and correlations). In this version of the paradigm, the cues represent measured credentials (e.g. test scores, grades, experience) of hypothetical applicants for clearly specified jobs or academic programs, and the criterion is an overall index of suitability (i.e. global ratings or decisions taken at face value or relative to a normative model). The principal focus of this type of task is the set of cognitive tendencies that people exhibit in aggregating numerical evidence into an overall judgment or prediction.

D. Threat evaluation. This is simply another version of the judgment task just described (C above). Formally it includes all the same properties. The only differences are the military "cover story" and the fact that it is more easily combined with formally different properties to constitute a plausible compound (or multiple-process) scenario. The next scenario is an illustration of just such a modification.

E. Two-stage judgment. The typical "policy-capturing" paradigm, as illustrated in (C) and (D), is a reasonable vehicle for studying how people deal with highly processed data in a highly structured judgment/decision setting. It focuses on the integration of processed, quantified cues -- a common requirement in modern systems. However, there are also many situations in which the input, or predictive evidence, does not come in processed form -- where judgment is based on

the raw observation of an event flow rather than a set of numbers. One can conceive of these latter problems as comprising two stages: a perceptual stage in which implicit "cue values" are generated from raw observations, and an integration phase in which the cues are aggregated (as in C and D above). Viewed in this way, the present scenario allows the subject to perform one or both judgments, separately or as a composite. The raw data are events (citings), observed over time, emanating from several distinct enemy positions. Rate of citings defines, within certain tolerances, a position's state of <u>readiness</u> for attack. That, together with its pre-defined <u>suitability</u> for attack, provides the cue values and importance weights necessary to formulate an aggregate threat evaluation over all enemy positions. In short, the "policy-capturing" problem of scenario D is combined with the frequency/probability estimation features of scenario A to yield the two-stage judgment problem. Of course, many variations can be used to satisfy particular experimental requirements (such as studying the effect of an <u>explicit</u> intermediate judgment on the ultimate quality of evaluation, as discussed below).

Together, these five scenarios have afforded us the opportunity to study many of the facets of judgment/decision behavior distinguished in the Howell & Burnett (1978) and the Hammond (1981) taxonomies, often within the same experiment. Use of each scenario in a number of pilot and formal studies has produced a fairly good understanding of its baseline performance, sensitivity, reliability, and other unique properties.

Most importantly, however, these scenarios have provided a wide range of partially overlapping task requirements within which to examine the generality

(and importance) of effects produced by particular kinds of manipulations. A common shortcoming in judgment/decision research, we believe, is that findings are limited to rather narrowly defined research paradigms (Bayesian inference, parameter estimation, confidence judgment, policy-capturing, gambling preference, optional stopping, heuristic judgment, belief perseveration, etc.). By contrast, one continuing objective of the research reported here has been to explain key variables, such as those suggested by Hammond (1980), Payne (1982) and our own previous work across paradigms. As illustrated in Figure 1, our five established task scenarios have begun to serve this function.

Independent variables. Most of the task variables investigated over the past three years are self-explanatory. As shown in Figure 1, they can be categorized for convenience roughly into four groups: those primarily involving information input, processing mode, response mode, and feedback. (Naturally we do not wish to imply that these functions are independent; only that the manipulations are focused at one or another point in the processing sequence). A few in the list which are not obvious are defined as follows.

4. stimulus enhancement refers to the addition of redundant information to particular events occurring on a display. It was introduced here in conjunction with efforts to induce availability effects, but has more general implications in that it represents a proposed means for combatting overload in computer-based systems (Knapp, Moses, & Gellman, 1982).
5. enhancement mode follows directly from (4). It refers to the particular way in which enhancement is implemented.

6. display order refers to sequential vs. simultaneous presentation of information.

7. display format has been manipulated primarily in terms of alphanumeric vs. graphic (or analog) coding. Both display variables were chosen primarily on the basis of their implications for the "cognitive continuum", but, of course, they represent common practical design options as well.

11. estimation requirement involves the insertion of explicit lower-order processing step(s), such as frequency estimation, in a higher-order judgment or decision task, such as diagnosis or choice.

12. response mode refers primarily to distinctions derived from the Howell & Burnett taxonomy, e.g. frequency estimation, probability estimation, prediction, cue integration (diagnosis), choice. These distinctions have been shown to influence the human's approach to uncertainty (Howell & Kerkar, 1981; 1982).

13 & 14. type and precision of feedback. These variables have typically been studied in the context of acquisition (e.g. MCPL). Here, however, our concern has been solely on asymptotic performance: the role of feedback characteristics in maintaining a level of performance that has already been established. (The entire project,with one exception, has had a performance rather than a learning focus).

CONCLUSION

While discussion of specific experiments and their results is reserved for the final section, a summary of the principal generalizations that either have emerged or seem to be emerging from this line of work is presented here. Since

these conclusions often cut across experiments, frequent reference is made to the domain defined in Figure 1.

1. Judgment is affected in both qualitative and quantitative ways by the manner in which information is displayed. That is, the human's strategies for weighting the importance of predictive information as well as the final product of those strategies are sensitive to the coding and ordering of that information. Graphic coding, for example, tends to promote more "holistic" processing than does numeric coding, and as a result, somewhat higher quality judgments (policy-capturing paradigm) as well as faster and better decisions, (optional-stopping paradigm). However, numeric coding is better suited to the more "analytic" task of learning an optimal cue weighting rule.

2. One can induce the availability heuristic by enhancing specific stimuli in an event stream. Consistent overestimation of enhanced relative to unenhanced emergencies was obtained in the emergency resource allocation problem (Scenario A), and the effect held up under replication. Surprisingly, the effect seems to be greatest at high (e.g. 8-10) and low (e.g. 1-2) objective frequencies, and all but disappears at frequency = 4. The reason for the resistance at this intermediate point, which was also replicated, is not clear. By varying where (early or late) and how often (once or on every occasion) in the event stream the enhancement occurred, it was found that the effect is not dependent on either -- all that mattered was whether at least one occurrence was enhanced. The schedule did, however, affect availability and retention of enhancement material.

3. <u>Time pressure does not operate in a simple fashion to reduce performance or to change mode of processing</u>. Time constraint is a variable that, according to Hammond's theory, should induce a more "intuitive" cognitive mode. Having incorporated it in at least one study under each scenario, however, we find that this variable does not have as uniform an effect as one might expect. In the optional stopping problem it combined with display format to produce what appeared to be a more intuitive approach. In policy-capturing studies, it seemed to reduce the subject's ability to apply his own rules. But in the two most complex scenarios (emergency resource allocation and two-stage judgment), its effects were less clear-cut. While not entirely unexpected, given the growing literature on attentional capacity characteristics (e.g. Wickens, 1980, 1984), this generalization points up the necessity for a more molecular approach to task description than is suggested by "continuum theory."

4. <u>Amount of information to be processed also exerts a complex influence on processing mode and performance</u>. It appears that requiring people to integrate more information items into a judgment does affect their processing approach, but <u>how</u> is dependent upon other factors (such as time constraint and display format). It is not simply a matter of inducing a more "holistic", intuitive form of judgment. This, of course, is not particularly surprising either in view of the huge literature on "load effects" and "capacity limitations" that has accumulated over the years (e.g. Moray, 1979, Navon & Gopher, 1979; Norman & Bobrow, 1975, Wickens, 1980; Wickens & Vidulich, 1982). However, much of this literature focuses on simpler,

less "cerebral", speeded (i.e. RT) tasks -- ones in which the rules relating stimuli and responses are well articulated or even "automated" (i.e. "rule-based" or "skill-based" tasks, to use Rasmussen's popular distinction, 1981). The present findings are among the relatively few bearing on the important question of how people handle progressively larger loads under various other task circumstances (i.e. those closer to "knowledge-based").

5. Manual pre-processing requirements can achieve results similar to those produced by automated pre-processing (aiding) in multi-stage judgment problems. If a human operator is required to estimate certain parameters of an ongoing event stream prior to making an overall judgment or decision based upon that evidence, the resulting judgment/decision is markedly improved. This is so even if the estimated values are not perfect. In one study, for example, threat diagnosis using such manual pre-processing compared very favorably with that using a "machine aid," even though the latter furnished 12-25% more accurate input (cue) values. Without either form of pre-processing threat evaluation dropped from about $r=.80$ to $r=.45$.

6. Asymptotic judgment performance based on explicit rules does not require feedback for maintenance; in fact, it deteriorates with outcome feedback. This was determined in one policy-capturing study where feedback type (process vs. outcome), precision, memory requirement, and information quality (cue-criterion relation) were examined. Findings showed that people are able to maintain the same level of consistency in applying a set of weighting rules without

process feedback as with it, but that outcome feedback -- even with a memory-aiding feature -- produces a progressive decrement.

7. The effects of rule availability and complexity are quite task dependent. According to Hammond's theory, having explicit rules available for integrating cues into judgments, particularly if they are not too complex, should encourage "analytic" processing. We found only weak evidence for this in the several studies where such manipulations were included. Of course, one would expect the value of such rules to be limited by other task features (e.g. speed/load stress), but even that relationship was not clearly established. Part of the difficulty lies in the fact that rule complexity is itself not easy to define. For example, the apparently simple rule of weighting cues equally turns out to be more complex subjectively than differential weighting.

II. Summary of Illustrative Experiment

The full details of the 15 experiments can be found in the quarterly progress reports and the publications cited in the last section below. The purpose of the present section is to illustrate, for each scenario, the kinds of studies undertaken and results obtained.

EMERGENCY RESOURCE ALLOCATION (SCENARIO A)

1. Frequency estimation and predictive choice as a function of qualitative event enhancement(#14, 15 in Figure 1). One of the more commonly cited "decision heuristics" is availability, the tendency for

people to judge event frequency or probability in terms of the ease with which instances of that event are brought to mind (Kahneman, Slovic, & Tversky, 1982; Tversky & Kahneman, 1974). Demonstrations of this phenomenon, however, have been limited principally to tasks involving pre-experimental knowledge for which availability differences have merely been assumed to exist.

In contrast, the purpose of the present studies was first, to determine whether differences in frequency judgment can be induced through enhancement of particular events during acquisition (i.e. an operation that should promote differential availability of the enhanced events); and secondly, to verify that such events are, indeed, differentially "available" at the time a frequency-based response is required.

The Emergency Resource scenario was a perfect vehicle for studying this issue. Certain events were enhanced during the dispatching phase of each problem (as the subject was acquiring familiarity with the frequentistic pattern of emergencies) by providing vivid descriptions of each actual emergency (e.g. details of an accident, fire etc.). Unenhanced events, acquired over the same time period, were merely identified as police or ambulance calls. A number of event categories at frequencies ranging from 2-16 per problem were selected for enhancement together with a like number of unenhanced control categories. As usual, subjects were required to make frequency judgments for both types of events during the test phase. In addition, they were presented with choice pairs pitting unenhanced and enhanced events of identical actual frequency against one another. The prediction, of course, was that both estimation and choice would favor the enhanced events. The design, therefore, was primarily a repeated measures model with actual frequency and enhancement as the within-subject variables.

The two experiments differed mainly in the between-groups manipulation that was crossed with the above variables. The first explored a response requirement variable that proved relatively unimportant. The second compared four different enhancement schedules: (a) only one instance of each designated event category enhanced in the early part of the problem; (b) one instance enhanced late in the problem; (c) all instances of the designated event category enhanced; and (d) all instances during the first of two sessions enhanced. The idea was to determine which operations produce the most frequency bias. In addition, retention measures were taken to determine whether the operations do, in fact, produce differences in availability.

Both studies yielded the same pattern of estimation (and choice) results: a significant tendency to overestimate the frequency of (or choose as more probable) the enhanced events relative to the unenhanced events, notably at high and low actual frequency levels. As illustrated in Figure 2 (Exp. 1) and 2a (Exp. 2), however,

----

Figures 2 and 2a about here.

----

there was a peculiar (and as yet unexplained) resistance to this bias at the middle level (frequency = 4 per session; represented as 8 and 12 in the two figures because two and three sessions were involved, respectively).

The manipulation of enhancement schedule did not produce reliable differences in the amount of bias (Figure 2), although it did yield significant differences in retention (hence availability). Substantially more events were available in memory, on the average, under continuously (vs.

singly) enhanced conditions. Thus it would appear that the availability of at least one "episode" in memory is sufficient to induce an inflated perception of event likelihood but that multiple (or "stronger") traces do not necessarily amplify the bias. In view of the fact that the schedules manipulation was intended only as a gross means of varying available information--not as a precise memory experiment--it would be improper to speculate on the detailed processes linking enhancement, event frequency, retention, and bias. Given the present findings, however, it would seem that a closer look at the microstructure of the "availability" heuristic would be desirable.

The principal message conveyed by these studies, then, is that DM does give undue weight to enhanced episodes when called upon to judge or use probabilistic information derived from direct observation of event occurrences. This could have important practical implications, for example, in the use of certain enhancement techniques in computer displays.

2. Judgment and decision performance as a function of time pressure and rule availability. Cognitive continuum theory suggests that DM is more likely to operate "analytically" (vs. "intuitively") to the extent that some processing rule is available and the DM has time to use it effectively. In keeping with this logic, a minor study was carried out to determine whether rule-based performance in a frequentistic task deteriorates to the level of a no-rule (intuitive) control under time pressure. A mixed model 2x2 design was used with time pressure and no pressure as the within-subjects variable, and presence or absence of rule-related instructions as the variable differentiating groups. The nature of the instruction manipulation concerned the manner in which emergency calls (i.e. the frequentistic events in this task) were generated. The rule-based

instructions emphasized the stochastic stationarity of the generating process and the importance of trying to learn and remember the frequency pattern. The _no-rule_ instructions left the DM to his own (intuitive) devices.

Contrary to expectations, rule-based performance held up better under time stress than did performance of the no-rule control group. That is, the expected superiority of rule-based judgment, which did seem to occur under stress-free conditions, did not diminish under time stress. If anything, it increased. This result raised a host of theoretical issues, some of which were explored using other task scenarios, but most of which were considered outside the scope of the current project. For present purposes, the main conclusion is that DM is not _necessarily_ forced into a qualitatively different processing mode as time pressure increases. The relative advantage to decision performance that is gained by simply _attending_ to the frequentistic aspects of a complex event scenario--an advantage that we have now demonstrated in a number of experiments--is not disturbed by time stress. Of course, an attention allocation "rule" is relatively simple; the same finding would not be expected for a more complex one. What is surprising is that so simple a rule does, itself, so consistently induce a different approach to the cognitive processing of frequentistic evidence.

OPTIONAL STOPPING (SCENARIO B)

Optional stopping performance under graphic and numeric CRT formatting (#10, 11 in Figure 1). Optional stopping tasks are among the most common of real-world decision problems but, due to their relative intractability for laboratory investigation, are among the least studied. One of the more consistent findings to emerge from the limited research on this generic problem

is oversampling, the tendency for DM to sample information (and delay terminal action) beyond an objectively defined optimum point (see, for example, Levine & Samet, 1973; Levine, Samet & Brahlek, 1975).

Continuum theory suggests that the analog or graphic display of information encourages a more holistic ("intuitive") mode of processing than does alpha-numeric formatting (which encourages a more serial, "analytical" approach). If so, one might expect graphic formatting to reduce the oversampling tendency in an optional stopping task, particularly if the DM were stressed at each decision point. Further, one would expect decision quality to be more resistent to time pressure under a graphic format due to its intuition-inducing properties.

Two studies were carried out to test this proposition. The first, using only self-paced presentation of information updates, was designed primarily to establish base level performance functions for the task and to estimate reasonable timing constraints for forced-pacing. The second study crossed time stress (three levels of forced-pacing) with format (graphic vs. numeric) in a mixed design, the format manipulation constituting the within-subjects variable. Three groups of 12 subjects each were thus defined according to stress level. One noteworthy feature of both experiments was that the displayed information was selected deliberately to be devoid of temporal trends (which, if present, would have favored a graphic display). Hence the test was extremely conservative.

As shown in Tables 1 and 2, the findings generally supported the hypotheses. That is, DM's did tend to sample fewer items prior to a terminal

----

Tables 1 and 2 about here.

----

decision under the graphic (analog) as opposed to the numeric format, and stress had a significantly greater (negative) impact on numeric than on graphic sampling and performance. In view of the very conservative nature of this test, one could only expect the pattern of differences to increase were trend information (which is typical of most real-world problems but naturally favors a graphic format) incorporated into the design. Thus, we consider the present findings rather substantial evidence for both the theoretical and practical implications of the Continuum Theory logic. Given time and an appropriate mental algorithm, numeric display can produce good optional stopping performance; the graphic mode, on the other hand, encourages a cognitive approach that is more resistant to time pressures.

The only finding that was not entirely consistent with predictions was a failure of numeric formatting to be clearly superior to the graphic mode under nonstressful conditions. This can be easily accounted for by specific task features. In addition, none of the conditions entirely eliminated the oversampling bias: in this regard, the studies are consistent with--and extend--findings from previous research.

INFORMATION INTEGRATION (SCENARIOS C & D)

1. Personnel selection performance as a function of load (number of cues), time stress, rule availability, and rule complexity, (#1 and 2 in Figure 1). The salient features of the standard information integration paradigm are (a) DM makes global judgments or predictions on the basis of multiple cues, (b) these judgments are regressed on the cue values to estimate the empirical regression weights for each DM, (c) these weights, taken as an index of the importance accorded each cue, are correlated with various other

"lens model" components to yield estimates of both overall judgment quality (e.g. "achievement" score) and the area in which judgment is deficient (e.g. "control" and "knowledge" scores).

Using this paradigm in the context of a familiar personnel-selection simulation, we manipulated variables from each of the major categories in Hammond's taxonomy (<u>structural</u>, <u>content</u>, and <u>presentation</u>) with the idea of identifying (a) the function by which overall judgment deteriorates, and (b) the principal cognitive source(s) of the decrement. Cognitive Continuum Theory suggests that increasing <u>time</u> pressure (a "presentation" variable), for example, should drive DM from an "analytic" to an "intuitive" mode. If so, there should be a shift both in the quality and nature of obtained judgments (i.e. in <u>achievement</u>, and some combination of <u>control</u> and <u>knowledge</u> scores). Similar effects should result from increases in the <u>cue load</u> (a "structure" variable) as well as the <u>availability</u> and <u>complexity</u> of an organizing principle ("content" variables).

Two studies were carried out using an identical mixed design with different combinations of variables assigned to the between and within-subjects dimensions. The first study crossed three levels of <u>cue load</u> (within-subjects) with two of <u>rule availability</u> (between-groups); the second, three levels of <u>time constraint</u> (within-subjects) with four of <u>rule complexity</u> (between-groups). Thus a total of six groups of 12 subjects each were used in the two studies.

The principal results are summarized in Tables 3 and 4. As expected, overall judgment quality declined as a function of cue load, time stress, and rule complexity over the range of each variable studied. More importantly,

---

Tables 3 and 4 about here.

---

however, the decrements appeared to operate through somewhat different cognitive mechanisms. Cue load had a particularly strong influence on DM's ability to formulate an appropriate weighting strategy (i.e. the knowledge component), even though the information necessary for doing so was explicitly provided. This is what one would expect if, in the language of Continuum Theory, the DM were to rely on an "intuitive" processing mode: knowing the proper rule had little effect since intuitive judgment is based on a simpler strategy. On the other hand, time constraint and/or weighting rule complexity seemed to have a greater impact on the application (i.e. the control component) than on the formulation of a proper weighting strategy. Performance broke down because DM had progressively greater difficulty carrying out his own preferred strategy with any consistency. Again, in Continuum Theory terminology, it was as though he were operating in a proper "analytic" mode but was unable to do all the required mental calculations in the time permitted.

The implication of the studies taken together is that not all task-induced "stressors" produce overall decrements in integrative judgment by the same means. Too much information may prompt a shift to an "intuitive" strategy; too little time may simply degrade one's consistent application of an "analytic" strategy. While not directly implied by Continuum Theory, this distinction could be regarded as a correllary with potentially important practical ramifications.

2. Display effects under simultaneous and sequential presentation of cues. (#7 in Figure 1). Following the same basic logic described above in conjunction with the optional-stopping scenario, it was hypothesized that a graphic display would induce a more holistic, intuitive approach to information integration than would a numeric display of the same cue values. Two studies were carried out to test this notion using very similar methodologies and designs. The first study crossed display mode with response requirement (judgment vs. choice) in a 2x2, within-subjects design. The second crossed display mode with cue load (4 vs. 6 cues), also in a 2x2, within-subjects design. The other principal difference was that the first study used simultaneous presentation of cues; the latter, sequential presentation (for reasons soon to be explained.)

Results of the first study were consistent with the hypotheses. As illustrated in Table 5, subjects tended to weight the cues more evenly

----

Table 5 & 6 about here.

----

under the graphic than under the numeric display: the weight accorded the intelligence cue (which was the most heavily weighted) was reliably smaller and those for motivation (moderately weighted) and experience (lightly weighted) reliably larger under the graphic display.

The second study sought to determine whether the obtained display effect was attributable to a more sequential (vs. holistic) processing strategy associated with the numerical format. On the assumption that serial presentation of cues under both formats would control this difference, the manipulation was replicated in the serial presentation mode. The results again

showed a significant cue x display interaction, but not nearly as pronounced or as systematic as in the first study (see Table 6). Individuals still were induced by the display to alter their weighting policy, but not necessarily in the direction of a more <u>even</u> weighting under the graphic format. Thus the serial control moderated, but did not eliminate, the display effect.

The safest conclusion to be drawn from these studies is that display format does affect the manner in which DM integrates cues just as it does optional stopping decisions. The differences are not entirely explained in terms of serial vs. holistic processing strategies, although graphic encoding does seem to promote a more holistic ("intuitive") approach. In this particular task, a holistic mode is advantageous in that it tends to offset the tendency to ignore lesser (but nonetheless potentially useful) predictive cues.

3. <u>Multiple-cue probability learning (MCPL) as a function of display format and weighting rule</u> (#9 in Figure 1). Unlike most of the research on this project, the present study addressed <u>acquisition</u> rather than asymptotic <u>performance</u> functions. Again, the question was whether display format affects cue weighting strategy, but in this case a <u>MCPL</u> paradigm was used. That is, a particular "environmental rule" was defined, and DM was required to learn what it was (or at least to develop his own weighting rule) based on feedback comparing his judgment to that of the optimal model. Attention was focused on <u>improvement</u> with reference to that model and on the various "lens model" measures that help to explain the cognitive basis for improvement.

The scenario in this case was the military threat evaluation problem (Scenario D) in which subjects judged overall threat of attack on one's own

position based on four types of intelligence data (cue values). The design was a 2x2 between-groups manipulation of display and optimal rule (equal vs. unequal weighting) conditions, with four trial blocks as a within-subjects variable. No explicit time constraint was used, and there was no reason to expect that subjects entered the task with any prior knowledge of--or differential weighting of--these particular cues. Because of these features and the learning orientation of the task, it was expected that subjects would adopt a more "analytic" than "intuitive" approach regardless of display, and if anything, the numerical display (which is more consistent with the analytical mode) would produce superior performance. Based on previous research, we expected an environmental model with unequal weights to prove easier to learn than one with equal weights.

The results confirmed these expectations. As shown in Table 7, the overall discrepancies between weights produced by DM and those of the optimal

----

Tables 7 and 8 about here.

----

model were significantly larger under the graphic format and the equal weighting rule. Substantial learning occurred for all groups, but particularly so under unequal weighting (Table 8). However, there was no format X blocks interaction. Thus, as expected, format affected performance (regardless of proficiency level attained); weighting rule affected acquisition. A noteworthy feature of both Table 7 and 8 data is that the absolute discrepancies were of a uniformly conservative nature (i.e. negative sign, suggesting failure to accord sufficient weight to the cues).

While an account of the various "lens model" measures is beyond the scope of this report, suffice it to say that they were consistent with the above generalizations in that improvement was greatest under the unequal weighting rule and display mode did not affect it. As shown in Table 9, this pattern was true for the overall correlations of DM's policy with the optimal

----

Table 9 about here.

----

rule (achievement score), the "knowledge" component (matching score, and the "control" component (optimality score). However, the "control" component contributed almost twice as much to the overall improvement as did the "knowledge" component. In other words, improvement comes about more in terms of learning to apply a rule consistently than in formulating that rule; and as we saw earlier, the rule that is applied tends to remain fairly conservative. This, of course, is all consistent with the view that acquisition is necessarily approached in an "analytic" fashion. "Intuitive" judgment is supposed to be relatively resistant to modification through experience (Hammond, 1981).

What this study shows, then, is that task conditions designed to favor "analytic" processing do, in fact, produce better performance under the display format (numeric) compatible with that mode. Moreover, acquisition--which is inherently "analytic" in the MCPL paradigm--is affected only by the difficulty of the weighting rule, not by display format. And finally, the overall tendency in this type of task is toward a conservative weighting policy.

TWO-STAGE JUDGMENT (SCENARIO E)

1. Diagnostic judgment as a function of manual and automated pre-processing of evidence and no pre-processing at all (#6 and 8 in Figure 1). In research on processes such as inference and integrative judgment, it is typical to present the subject with highly processed data (e.g. cue values, diagnostic impact values, etc.). The same is true, of course, in highly sophisticated decision systems (e.g. where various decision aiding algorithms are used). However, most real-world decisions are still made on the basis of relatively unprocessed evidence--often, raw observations made by DM over time.

The purpose of this line of research, therefore, was to determine how the quality of human judgment (in this case, military threat diagnosis) is affected by various levels of pre-processing applied to the raw predictive events when such processing is done manually and through aiding. In essence the paradigms extended the standard "policy-capturing" approach to a situation in which cue values (processed predictors) were derived from a more fundamental set of events (raw observations)by man, machine, or a combination.

Two studies involved between-groups comparison of overall threat judgments made under conditions in which overt estimates of observed activity was required (estimation groups) or was not required (no-estimation groups) for identical sets of raw observations (enemy citings). Activity level for various regions constituted the basis for environmental (i.e. "true") threat. Thus either estimated or actual activity levels could be used as cues in an information-integration paradigm (as in Scenario D); or subjects could be required to make estimates directly from raw observations.

The first study was primarily concerned with the question of whether requiring an overt estimate of activity level (cue values) enhances threat

judgments; hence the <u>no-estimation</u> vs. <u>estimation</u> group comparison was of chief interest. The second study included a third treatment group for which cue values (activity levels) were <u>computed automatically</u> and presented in the standard numerical form (as in Scenario D). Each study also included other conditions which, for purposes of clarity, will not be discussed here.

The main findings are expressed in terms of two measures: quality of the overall threat assessments (correlations of judged with optimal values), and distribution of importance weights across enemy regions (empirical and optimal B-weights). As shown in Tables 10 and 11 the <u>estimation</u> requirement

---

Table 10 and 11 about here.

---

clearly enhanced overall assessment performance in the absence of any aiding or under minimal aiding (simple tabulation) conditions. It did not, of course, when the cue values were actually <u>computed</u>. What is particularly noteworthy, however, is that mere estimation of cues produced threat evaluations ($r=.79$) on a virtual par with those obtained when the cues were actually computed ($r=.83$). This happened despite the fact that the cue estimates were far from perfect (as low as 76% accuracy under some conditions). Without the benefit of estimation or aiding on the other hand, threat evaluation dropped to $r=.47$. Obviously, people are not very adept at making global judgments directly from raw observations, even when the observed evidence is quite simple.

These conclusions are supported further by the distribution index (see Tables 12 and 13) which shows that improvement in overall performance

corresponds to a better distribution of importance weights across cues under the computation and estimation conditions.

Tables 12 and 13 about here.

In general, then, the results of these studies suggest that both the estimation requirement and aiding serve to cast the predictive information into a form conducive to integration (increasing, in a sense, its compatibility with the required cognitive operations). Such pre-processing presumably simplifies the ultimate integration step, but in a way that encourages preserving rather than discarding predictive information. Without such an explicit pre-processing step, DM tends to simplify in other, less predictive ways (e.g. overselection.

III. Reports Generated Under the Project

Howell, W. C., & Kerkar, S. P. A test of task influences in uncertainty measurement. Organizational Behavior and Human Performance, 1982, 30, 365-390.

Fontenelle, G. The effects of task characteristics on the availability heuristic for judgments under uncertainty. Technical Report No. 83-1, May, 1983.

Kerkar, S. P. A critical analysis of the uses of multiple regression in the study of human judgment. Technical Report No. 83-2, July, 1983.

Goldsberry, B. S. In search of the components of task induced decrements. Technical Report No. 83-3, August, 1983.

Schwartz, D. R., & Howell, W. C. Optional stopping performance under graphic and numeric CRT formatting. Technical Report No. 84-1, June, 1984.
(A somewhat modified version under review by Human Factors).

Kerkar, S. P., & Howell, W. C. The effect of information display format on multiple-cue judgment. Technical Report No. 84-2, June, 1984.

Goldsberry, B. S. The effect of feedback and predictability on human judgment. Technical Report No. 84-3, August, 1984.

Friedman, L., Howell, W. C., & Jensen, C. R. Diagnostic judgment as a function of the pre-processing of evidence. Technical Report No. 84-4, October, 1984.

(A somewhat modified version under review by Human Factors).

Fontenelle, G. & Howell, W. C. A replication and extension of the inducement of the availability heuristic. Technical Report No. 84-5, December, 1984.

Fontenelle, G. & Howell, W. C. On the inducement and verification of availability bias in a simulated dispatching task. (In preparation for submission to an archival journal).

References

Brunswik, E. Perception and the representative design of psychological experiments. Berkeley, CA: University of California Press, 1956.

Dawes, R. M., & Corrigan, B. Linear models in decision making. Psychological Bulletin, 1974, 81, 95-106.

Goldberg, L. R. Man vs. model of man: A rationale, plus some evidence, for a method of improving on clinical inferences. Psychological Bulletin, 1970, 73, 422-432.

Hammond, K. R. Principles of organization in intuitive and anlytical cognition. Report No. 231. University of Colorado, Center for Research on Judgment and Policy, February, 1981.

Hammond, K.R. The investigation of research in judgment and decision making. (Technical Report 226). Boulder, Colorado: University of Colorado, Center for Research in Judgment and Policy, 1980.

Hammond, K. R., & Summers, D. A. Cognitive control. Psychological Review, 1972, 79, 58-67.

Howell, W.C., & Burnett, S.A. Uncertainty measurement: A cognitive taxonomy. Organizational Behavior and Human Performance, 1978, 22, 45-68.

Howell, W. C. & Kerkar, S. P. Uncertainty measurement in a complex task as a function of response mode and event type characteristics. Technical Report No. 81-1, Department of Psychology, Rice University Research Report Series, February, 1981.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.) Judgment under uncertainty: Heuristics and biases. New York: Cambridge, 1982.

Knapp, B. G., Moses, F. L., & Gellman, L. H. Information highlighting on complex displays. In A. Badre & B. Schneiderman (Eds.).

Moray, N. (Ed.) Mental Workload: Its Theory and Measurement. New York: Plenum

Navon, D., & Gopher, D. On the economy of the human processing system. Psychological Review, 1979, 86, 254-255.

Norman, D., & Bobrow, D. On data-limited and resource-limited processing. Journal of Cognitive Psychology, 1975, 7, 44-60.

Payne, J. W. Contingent decision behavior. fPsychological Bulletin, 1982, 92, 382-402.

Rasmussen, J. Models of mental strategies in process control. In J. Rasmussen & W. Rouse (Eds.), Human detection and diagnosis of system failures. New York: Plenum, 1981.

Sage, A. P. Behavioral and organizational considerations in the design of information systems and processes for planning and decision support. IEEE Transactions, 1981,SMC-11, No. 9, 640-678.

Slovic, P. Toward understanding and improving decisions. In W. C. Howell & E. A. Fleishman (Eds.), Human Performance and Productivity, Vol. 2, Human Information Processing and Decision Making. Hillsdale, N. J.: Erlbaum, 1981.

Slovic, P. Fischhoff, B., & Lichtenstein, S. Behavioral decision theory. Annual Review of Psychology, 1977, 28, 1-39.

Tversky, A., & Kahneman, D. Judgment under uncertainty: heuristics and biases. Science, 1974, 185, 1124-1131.

Wickens, C.D. Engineering psychology and human performances. Columbus, Ohio: Merrill, 1984.

Wickens, C. D., & Vidulich, M. S-C-R compatibility and dual-task performance in two complex information processing tasks:

Threat evaluation and fault diagnosis. (EPL-82-3/ONR-82-3). Champaign, IL: University of Illinois, Engineering-Psychology Laboratory, December, 1982.

Wickens, C. D. The structure of attentional resources. in R. Nickerson and R. Pew (Eds.), Attention and Performance VIII. Hillsdale, N.J.: Erlbaum, 1980.

Wohl, J. G. Force management decision requirements for Air Force Tactical Command and Control. IEEE Transactions, 1981, SMC-11, No. 9, 618-639.

Table 1

Summary of Major Sampling Effects Expressed as the Number of Updates Prior to a Terminal Decision (Experiment 2)

Display Format	Trial Block	Stress Level (Updates/sec) .6	1.4	3.3	Mean
	1	4.72	5.04	5.00	4.92
Analog	2	4.66	5.05	5.01	4.91
	3	4.86	5.00	4.88	4.91
	Mean	4.75	5.03	4.96	4.91
	1	5.17	4.70	4.87	4.91
Numeric	2	5.49	4.91	5.09	5.16
	3	5.51	4.72	5.30	5.18
	Mean	5.39	4.78	5.08	5.08
Mean		5.07	4.90	5.02	5.00

(1) Format x Stress $F(2,30) = 4.38$, $p = 0.021$

(2) F x Trials $F(2,60) = 3.58$, $p = 0.034$

Table 2

Mean Accuracy Scores (% correct) for the Various Display Format

and Time Stress Conditions in Experiment 2

Display Format	Order	Stress Level (Updates/sec) .6	1.4	3.3	Mean
	AN	60	52	47	53
Analog	NA	56	53	48	52
	Mean	58	53	47	53
	AN	58	51	43	51
Numeric	NA	55	38	41	45
	Mean	57	45	42	48
Mean		57	49	45	50

(1) Format $F(1,30) = 15.47$, $p < 0.001$

(2) Stress $F(2,30) = 11.84$, $p < 0.001$

(3) F x S $F(2,30) = 2.11$, $p = 1.39$ (ns)

(4) F x S x Order $F(2,30) = 3.00$, $p = 0.064$ (ns)

Table 3

Mean Scores Obtained under the Six Experimental Conditions in Experiment 1 on All Five Measures

Experimental Condition	Product: Hit Rate(%)	Product: Ach.(r)	Process (r): Knowledge(G)	Process (r): Control(C)	Process (r): (GxC)	Relia-bility(R)
Strategy Not Available (Group I)						
3 cues	49	.93	.98	.95	.93	.92
4 cues	38	.82	.89	.92	.82	.81
5 cues	36	.75	.83	.92	.76	.88
collapsed over quantity	41	.83	.90	.93	.84	.87
Strategy Available (Group II)						
3 cues	59	.89	.97	.92	.89	.87
4 cues	39	.79	.92	.85	.78	.85
5 cues	39	.73	.85	.85	.72	.77
collapsed over quantity	46	.80	.91	.87	.80	.83
Collapsed Over Groups						
3 cues	54	.91	.97	.94	.91	.90
4 cues	39	.80	.91	.89	.81	.83
5 cues	38	.74	.84	.89	.75	.83
collapsed over quantity	44	.82	.91	.91	.82	.85

Table 4

Mean Scores Obtained under the Twelve Experimental Conditions in Experiment 2 on All Five Measures

Experimental Condition	Measures: Product: Hit Rate(%)	Product: Ach(r)	Process(r): Knowledge(G)	Process(r): Control(C)	Process(r): (GxC)	Process(r): Reliability(R)
4 Equal Wgt. Strategy (Group I)						
15 sec.	74	.94	.98	.96	.94	.94
10 sec.	71	.89	.98	.91	.89	.83
5 sec.	54	.83	.97	.85	.82	.72
collapsed over times	66	.89	.98	.91	.88	.83
3 Equal Wgt. Strategy (Group II)						
15 sec.	78	.95	.99	.96	.95	.93
10 sec.	70	.94	.98	.96	.94	.93
5 sec.	56	.91	.99	.92	.91	.86
collapsed over times	68	.93	.99	.95	.93	.91
2 Equal Wgt. Strategy (Group III)						
15 sec.	49	.91	.98	.93	.91	.88
10 sec.	48	.86	.97	.89	.86	.81
5 sec.	36	.87	.96	.87	.84	.79
collapsed over times	44	.87	.97	.90	.87	.83
0 Equal Wgt. Strategy (Group IV)						
15 sec.	44	.88	.98	.90	.88	.82
10 sec.	43	.88	.95	.93	.88	.90
5 sec.	41	.84	.97	.87	.87	.83
collapsed over times	43	.88	.97	.90	.88	.85

Table 5

Mean Raw Score Regression Weights (adjusted)) for the Four Cues Under the Numerical and Graphic Formats (Experiment 1)

Display Format	Cues: Intelligence	Motivation	Skill	Experience
Numerical	1.12	.57	.37	.03
Graphic	.97	.70	.33	.15

Note: Adjusted regression weights were obtained by multiplying the raw score weights and standard deviations of cue values to equate scale differences among the four cues.

Table 6

Mean Raw Score Weights for the Four Cues Under the Numerical and Graphic Formats (Experiment 2)

Display Format	Intelligence	Motivation	Skill	Experience
Numerical	.82	.73	.46	.78
Graphic	.75	.94	.57	.93

Note: Since the scales had equal SDs, no adjustment was necessary for the analyses. The means in the Table are, however, multiplied by the SDs to make them comparable to those of Experiment 1.

Table 7

Total Deviations of Raw Scores Regression Weights (summed across all four cues) Derived from Subjects' Policies and the Optimum Policies for the Four Experimental Groups

Group	Signed Deviation	Unsigned Deviation
Numeric Unequal	-4.48	7.29
Graphic Unequal	-7.69	9.39
Numeric Equal	-8.82	9.14
Graphic Equal	-11.36	11.49

Table 8

Deviation Scores Averaged Over Display Formats for the Four Trial Blocks

	Block			
Weighting Rule	1	2	3	4
Equal	-11.26	-11.75	-9.73	-7.60
Unequal	- 8.10	- 5.45	-5.40	-5.39

Table 9

Summary of Correlational Measures Used to Index Subjects' Performance

Performance Index	Optimal Weighting Rule	Learning Block 1	2	3	4
Achievement (Overall)	Unequal	.47	.61	.68	.69
	Equal	.47	.45	.55	.50
	X	.47	.53	.61	.59
Matching Coefficient (Knowledge)	Unequal	.80	.87	.89	.92
	Equal	.79	.71	.79	.86
	X	.79	.79	.84	.89
Optimality Coefficient (Control)	Unequal	.51	.68	.73	.73
	Equal	.51	.46	.57	.61
	X	.51	.57	.65	.67

Note: The group means are collapsed across the numerical and graphical display formats.

Table 10

Mean Correlations Between Actual and Optimal Threat Assessments for Experiment 1

	Aiding			
	Automatic Tabulation		Self Tabulation	
Group	M	SD	M	SD
Estimation	.82	.09	.82	.15
No Estimation	.75	.21	.74	.12

Table 11

Mean Correlations Between Actual and Optimal Threat Assessments for Experiment 2

	Aiding					
	None		Tabulation		Computation	
Group	M	SD	M	SD	M	SD
Estimation	.79	.08	.86	.09	.83	.07
No Estimation	.47	.18	.77	.10	.90	.07

Table 12

Mean B-Weights of Each Region for Various Experiment 1 Conditions

	Region							
	1		2		3		4	
Group	M	SD	M	SD	M	SD	M	SD
Est.	5.41	1.31	3.14	1.24	1.48	1.60	0.82	1.07
No Est.	5.00	1.72	2.40	1.24	0.87	1.54	0.66	1.61
Optimal	4.00		3.00		2.00		1.00	

Table 13

Mean B-Weights For Each Region For the Various Experiment 2 Conditions

Unaided Group

	Region							
	1		2		3		4	
Group	M	SD	M	SD	M	SD	M	SD
Est.	5.51	1.42	2.65	1.50	0.73	1.08	0.31	1.36
No. Est.	4.12	2.03	1.52	1.80	-.07	2.09	-.88	1.74
Optimal	4.00		3.00		2.00		1.00	

Table 13, cont'd

Mean B-Weights For Each Region For the Various Experiment 2 Conditions

Tabulation Group

	Region							
	1		2		3		4	
Group	M	SD	M	SD	M	SD	M	SD
Est.	5.48	2.21	3.12	1.18	1.20	1.61	0.38	1.20
No Est.	4.67	2.06	2.70	1.02	1.14	1.40	0.05	2.39
Optimal	4.00		3.00		2.00		1.00	

Computation Group

	Region							
	1		2		3		4	
Group	M	SD	M	SD	M	SD	M	SD
Est.	4.64	1.34	3.16	1.01	1.00	1.07	0.78	1.39
No Est.	6.01	1.31	3.51	1.26	1.01	0.95	0.33	0.88
Optimal	4.00		3.00		2.00		1.00	

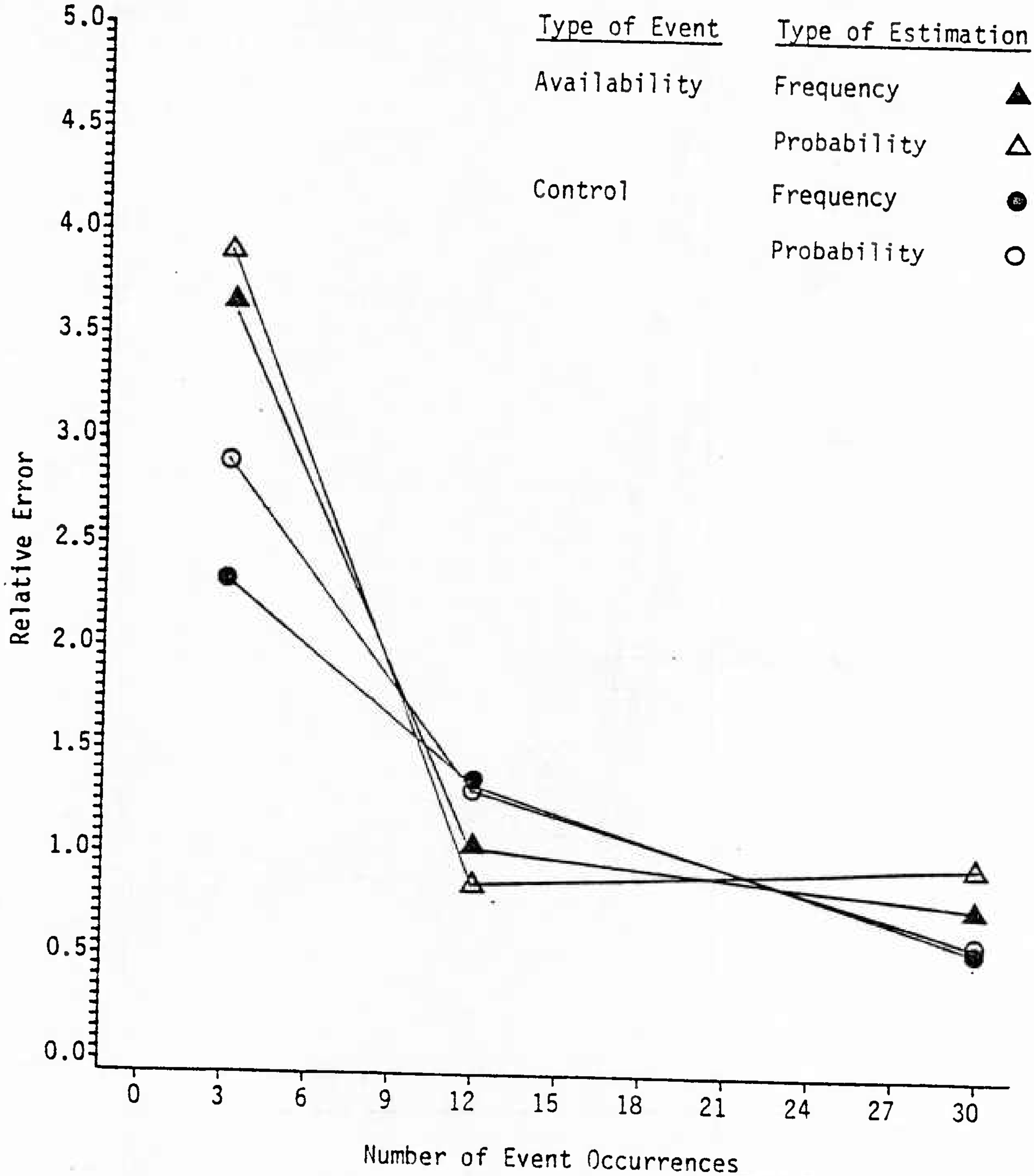

Figure 2. Mean error scores for the various frequency levels and enhancement schedules.

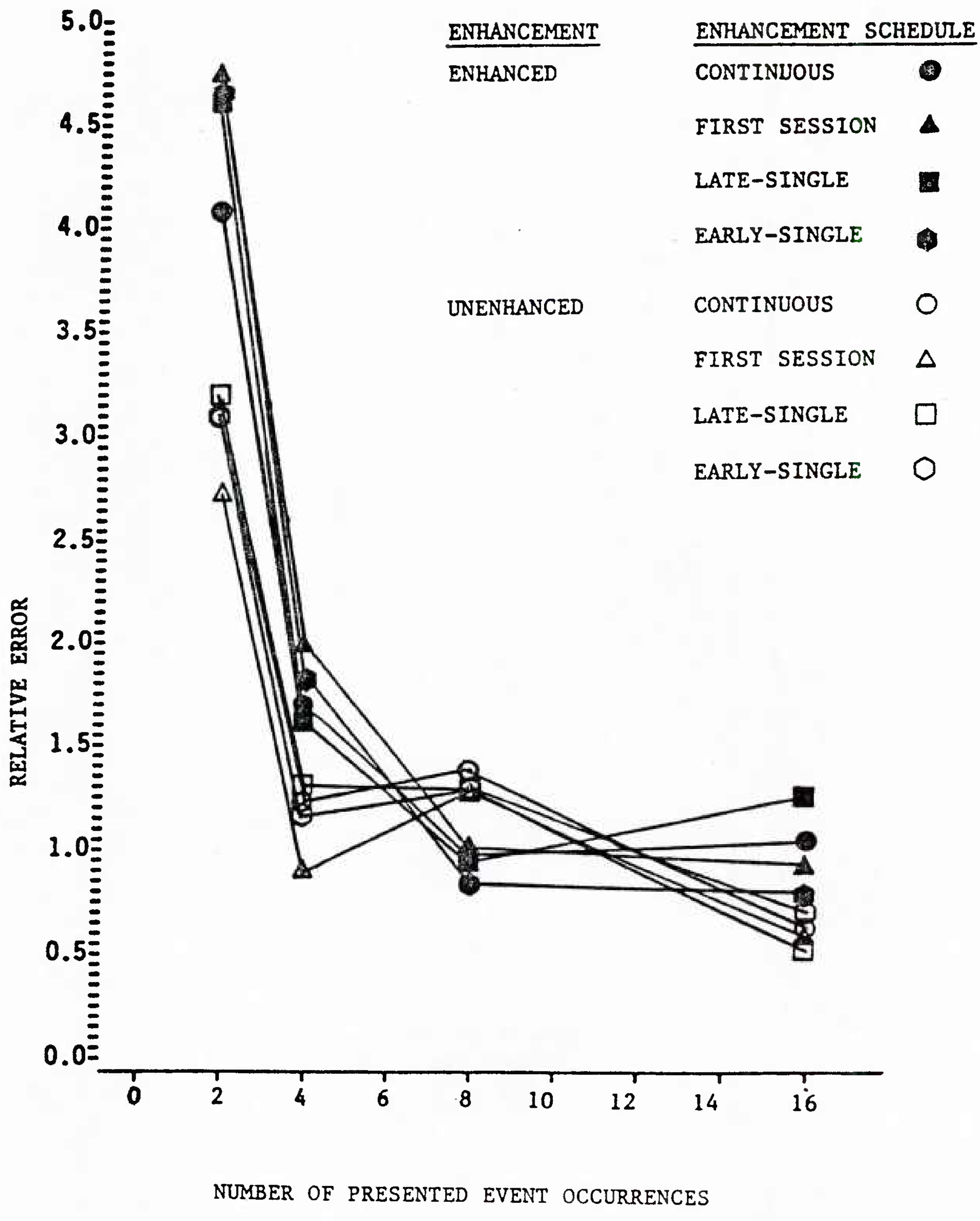

Figure 2a. Mean error scores for the various frequency levels and enhancement schedules.